العنوان: A Proposed Statistical Approach For Testing in Vivo Bioequivalence
المصدر: المجلة المصرية للدراسات التجارية
الناشر: جامعة المنصورة - كلية التجارة
المؤلف الرئيسي: Ibrahim, Ibrahim Hassan
المجلد/العدد: مج23, ع2
محكمة: نعم
التاريخ الميلادي: 1999
الصفحات: 7 - 25
رقم MD: 659357
نوع المحتوى: بحوث ومقالات
قواعد المعلومات: EcoLink
مواضيع: التحليل الإحصائي، المقاربات الإحصائية، الإحصاء الرياضي
رابط: http://search.mandumah.com/Record/659357

A Proposed Statistical Approach For Testing In Vivo Bioequivalence

*Ibrahim Hassan Ibrahim**

ABSTRACT

The basic goal of bioequivalence studies is to provide statistical evidence of bioequivalence between a reference and a test drug . From reviewing literature, it has been noticed that all of the bioequivalence studies depend on some univariate statistical analysis only . This type of analysis could be misleading and might lead to incorrect conclusions, because the univariate analysis neglects the intercorrelation between dependent variables This paper suggested the use of the multivariate statistical analysis in addition to the univariate analysis to demonstrate bioequivalence, that is because the multivariate analysis takes advantage of the data covariance structure, which is important, especially when two responses are correlated . A numerical example is provided to illustrate the basic idea of this paper .

KEY WORDS : *Crossover design; Nested model; ANOVA; MANOVA .*

1. INTRODUCTION

Bioequivalence studies have received much attention in the recent literature . The basic concern of bioequivalence studies is to test if a reference drug and a test drug are equivalent, in other word, testing if there is no significant difference between the reference and the test drug . These tests are based upon some pharmacokinetic parameters, such as area under the plasma or blood concentration-time curve (AUC) and peak drug concentration (C_{max}) and the time to peak drug concentration (T_{max}) . If there is no significant difference between the reference and the test drug, we can say that these two drugs are bioequivalent .

In Egypt, bioequivalence researches have received much attention in the last few years because of their importance in improving the Egyptian industry of medical drugs . There are centers for bioavailability studies in most of the Egyptian Universities located in Faculties of Pharmacy such as in Cairo University, Tanta University and Alexandria University . These centers are responsible for testing bioequivalence between a reference drug, usually manufactured by a foreign company and a test drug manufactured by a local Egyptian company . Foreign currency can be saved in case of bioequivalence between the imported (reference) drug and the locally produced (test) drug, because the test drug will be produced and sold on a national

**Assoc. Prof., Dept. of Mathematics & Statistics, Faculty of Commerce, Tanta University .*

scale . Despite the important role of these centers, it has been noticed that most of these centers apply simple univariate statistical analysis and even some tests such as the t-test or the Wilcoxon test to prove bioequivalence . This type of analysis is misleading and will lead to wrong conclusions . This paper suggested the use of multivariate statistical analysis in addition to the univariate analysis to demonstrate bioequivalence between a reference drug and a test drug, that is because the multivariate analysis takes advantage of the data covariance structure, which is very important especially when there is a correlation between two dependent variables .

The main goal of this research is to provide an alternative statistical approach that can be used to prove statistical bioequivalence between reference and test drug . This approach depends on testing the differences between test and reference drugs in a univariate scale as well as in a multivariate scale .

2. RESEARCH TERMINOLOGY

Definitions of some terminology that are used in this paper are provided according to " Guidance for Industry " published by Center for Drug Evaluation and Research (CDER), U.S. Department of Health and Human Services, Food and Drug Administration (FDA), which can be found on the following Internet web site http://www.fda.gov.cder/guidance/index.html. :

1. Pharmacokinetic parameters : The bioavailability (BA) measures used to test for bioequivalence (BE) such as :

(i) AUC_{0-t} : Area under the plasma/blood concentration-time curve from time zero to time t, where t is the last time point with measurable concentration .

(ii) $AUC_{0-\infty}$: Area under the plasma/blood concentration-time curve from time zero to time infinity, where $AUC_{0-\infty} = AUC_t + C_t / \lambda_Z$, C_t is the last measurable drug concentration and λ_Z is the terminal or elimination rate constant calculated according to an appropriate method.

(iii) C_{max} : Peak drug concentration .

(iv) T_{max} : The time to peak drug concentration obtained directly from the data without interpolation.

2. ***In vivo*** **and** ***In virto*** **:** *in vivo* means testing done on humans, while *in virto* means testing the physicochemical properties of the drug such as hardness, weight and dissolution . This paper will be limited to in vivo testing only .

3. REVIEW OF STATISTICAL LITERATURE

Schuirmann, (1987) compared various methodologies for assessing drug product equivalence and described the confidence interval approach that should be applied to (normally distributed) data from a two-period crossover design to the individual parameters of interest (e.g., AUC and C_{max}).

In 1992 the Center for Drug Evaluation and Research (CDER), FDA, USA issued a guidance recommending that a standard in vivo bioequivalence study design should be based on administration of the test and reference products on separate occasions to healthy subjects, either in single or multiple doses, with random assignment to the two possible sequences of drug product administration . The guidance stated that samples of plasma or blood should be analyzed for drug concentrations, and pharmacokinetic parameters be obtained from the resulting concentration-time curves. The guidance suggested that the pharmacokinetic parameters be analyzed statistically to determine if the test and reference products yielded comparable values. Statistical analysis for pharmacokinetic parameters, such as area under the curve (AUC) and peak concentration (C_{max}), was based on a test procedure termed "the two one-sided tests procedure," which determined whether the average values for pharmacokinetic parameters measured after administration of the test and reference products were comparable (i.e., average bioequivalence) . This recommended procedure involved the calculation of a 90% confidence interval for the ratio of the averages of the test and reference product . To establish bioequivalence, the calculated confidence interval was to fall within a bioequivalence limit, usually 80-125% for the ratio of the product averages.

In 1997 CDER updated the 1992 guidance, and recommended that the average bioequivalence method for determining bioequivalence should be replaced by two new approaches, termed *population* and *individual bioequivalence*. The population bioequivalence approach assesses the total variability of the metric in the population. The individual bioequivalence approach assesses the within-subject variability as well as the subject-by-formulation interaction . The population and individual bioequivalence approaches reflect differences in the objectives of bioequivalence testing at various stages of drug development . These differences are explained in the concepts of *prescribability* and *witchability* . *Prescribability* refers to the clinical setting where a practitioner prescribes a drug product to a patient for the first time. In

this case, the prescriber relies on an understanding that the average performance of the drug product has been well characterized and relates in some definable way to the clinical trial material on which safety and efficacy data were generated . *Switchability* refers to the setting where a practitioner transfers a patient from one drug product to another . The CDER recommended that a crossover design, such as the standard two-formulation, two-period, two-sequence crossover design, may be used to generate data for assessment of population bioequivalence. While for the individual bioequivalence, the guidance recommended the use of the within-subject variances for the test and reference metric and subject-by-formulation interaction variance component . A replicated-crossover design of the bioequivalence study should be used to estimate these parameters. The following four-period, two-sequence, two-formulation design is recommended:

		Period			
		1	**2**	**3**	**4**
Sequence	**1**	T	R	T	R
	2	R	T	R	T

where:

T = Test Formulation

R = Reference Formulation

For this design, the same lots of the test and reference formulations should be used for the replicate administration. Each period should be separated by an adequate washout period. The 1992 guidance, recommended also some other crossover designs such as the three-period design, as fellows :

		Period		
		1	**2**	**3**
Sequence	**1**	T	R	T
	2	R	T	R

To achieve the same statistical power to conclude bioequivalence, a greater number of subjects would be needed for the three-period design compared to the recommended four-period design . The guidance also, recommended the use of a replicated-crossover design with no more than two sequences to avoid ambiguities in the estimation of the parameters in the bioequivalence criterion .

Liu, Jen-Pei, (1998) discussed the Advantages and drawbacks of the current statistical procedures for assessment of individual bioequivalence with emphasis on the aggregate-based criteria . He showed that drug switchability requires the evidence of individual bioequivalence which refers to the comparison of the closeness between the two distributions of the pharmacokinetic (PK) responses from the same subject obtained under the repeated administrations of the test and reference formulations. Liu proposed an intersection-union test based on disaggregate criteria for the evaluation of individual bioequivalence.

In 1999 CDER updated the 1997 guidance, and suggested using three criteria for bioequivalence (BE) : average, population, and individual criteria . The average BE approach focuses only on the comparison of population averages of a BE measure of interest and not on the variances of the measure for T and R products. The average BE method does not assess a subject-by-formulation interaction variance, that is, the variation in the average T and R difference among individuals . In contrast, population and individual BE approaches include comparisons of both the averages and variances of the study measure. The population BE approach assesses the total variability of the measure in the population. The individual BE approach assesses within-subject variability for T and R product, as well as the subject-by-formulation interaction . The recommended statistical model is a mixed-effects or two-stage linear model based upon the logarithmic transformation of the BA measures (e.g., AUC and C_{max}) . Each subject, j, theoretically provides a mean for the log-transformed BA measure for each formulation, μ_{Tj} and μ_{Rj} for the T and R formulations, respectively. The model assumes that these subject-specific means come from a distribution with population means $\hat{\mu}_T$ and $\hat{\mu}_R$ and between-subject variances σ^2_{BT} and σ^2_{BR} , respectively. The model allows for a correlation, ρ, between $\hat{\mu}_{Tj}$ and $\hat{\mu}_{Rj}$. The subject-by-formulation interaction variance component, σ^2_{ρ}, is related to these parameters as follows:

$$\sigma^2 = \text{variance of } (\hat{\mu}_{Tj} - \hat{\mu}_{Rj})$$

$$= (\sigma_{BT} - \sigma_{BR})^2 + 2(1-\rho)\,\sigma_{BT}\,\sigma_{BR}$$

For a given subject, the observed data for the log-transformed BA measure are assumed to be independent observations from distributions with means μ_{Tj} and μ_{Rj}

and within-subject variances σ^2_{WT} and σ^2_{WR}. The total variances for each formulation are defined as the sum of the within- and between-subject components (i.e., $\sigma^2_{TT} = \sigma^2_{WT} + \sigma^2_{BT}$ and $\sigma^2_{TR} = \sigma^2_{WR} + \sigma^2_{BR}$). For analysis of crossover studies, the means are given additional structure by the inclusion of period and sequence effect terms. The general structure of a BE criterion is that a function (γ_1) of population parameters should be demonstrated to be no greater than a specified value (γ_2). This is accomplished by testing the hypothesis $H_0 : \gamma_1 > \gamma_2$ versus $H_A : \gamma_1 \leq \gamma_2$ at a desired level of significance . Rejection of the null hypothesis means that the estimate of γ_1 is statistically significantly less than γ_2, results in a conclusion of BE . The choice of γ_1 and γ_2 differs in average, population, and individual BE criteria . The normal-theory is recommended for the analysis of log-transformed BE measures . For average BE, the general approach is to construct a 90% confidence interval for the quantity (μ_T - μ_R) to reach a conclusion of average BE . Due to the nature of normal-theory confidence intervals, this is equivalent to carrying out two one-sided tests of hypothesis at the 5% level of significance (Schuirmann 1987) . For a two-treatment, two-period, two-sequence (2 x 2) randomized crossover design, the statistical model includes factors accounting for the following sources of variation: sequence, subjects nested in sequences, period, and treatment .

Wang, Weizhen, (1999) proposed an exact level-α test for testing individual bioequivalence under normality assumption . Wang proved that a (2 x 3) crossover design is sufficient for assessing individual bioequivalence and there is no need for higher-order crossover designs .

Friesen, M. and Walker, S., (1999) used computer simulation to determine the likelihood of two bioequivalent (vs. reference) generic warfarin formulations (with varying bioavailability) passing current bioequivalence criteria against each other at varying bioavailability. They proved that there exists a possibility for two generic products, which are known to be bioequivalent to a reference formulation, to be bio-in-equivalent to each other . Therefore they propose a 5% limit in the T/R mean ratio, be added to the criteria for evaluation of narrow therapeutic index drugs.

4. THE PROPOSED STATISTICAL APPROACH

In the literature on bioequivalence assessment, there has been a common factor and main understanding from the very first that the proper statistical analysis of a standard comparative bioavailability study requires univariate methods only . In the present paper, the multivariate methods as well as the univariate methods are used . That is because, in a two-period crossover trial, any difference between the sequence groups in either one of the study periods is at variance with identity of the drug formulations under comparison . Consequently, for purposes of establishing the equivalence of both treatments, it is not sufficient to look at the direct treatment effects only using the ANOVA technique . It is also important to use the MANOVA technique which takes advantage of the data covariance structure to simultaneously test the equality of means from different responses .

4.1 The Univariate Analysis of Variance ANOVA :

The CDER in 1999 recommended the use of a (2X4) crossover design while, Wang,Weizhen, (1999) has proved that a (2X3) crossover design is sufficient for testing bioequivalence . The proposed approach is to test bioequivalence by using the ANOVA method, by a (2X2) mixed effects model . From practical point of view this model is more appropriate, because obtaining the pharmacokinetic parameters require taking blood samples from the volunteers every half hour for a two day period, two times in a (2X2) design and four times in a (2X4) design which is too much for the volunteers . This paper suggests the use of a two-treatment, two-period, two-sequence randomized crossover design as follows :

		Period	
		1	2
	1	R	T
Sequence			
	2	T	R

The statistical model is :

$$Y_{ijkl} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + \delta_l + \varepsilon_{ijkl} \qquad \text{.....(1)}$$

Where,

Y_{ijkl} : the response (such as AUC or C_{max} or T_{max}) of treatment "l" in period "k" for subject "j" in sequence "i" .

α_i : the sequence (group) effect, i = 1, 2 .

$\beta_{j(i)}$: the subjects j nested within sequences i , for j = 1, 2, 3, ..., n .

γ_k : the period effect, k = 1, 2 .

δ_l : the treatment effect, l = 1 for referense drug R, and l = 2 for test drug T .

ε_{ijkl} : the random error of the model .

Assumptions of the model :

(1) $\alpha_1 + \alpha_2 = 0, \quad \gamma_1 + \gamma_2 = 0, \quad \delta_1 + \delta_2 = 0$

(2) $\varepsilon_{ijkl} \sim$ i.i.d. $N(0, \sigma^2_\varepsilon)$, $\beta_{j(i)} \sim N(0, \sigma^2_\beta)$

(3) ε_{ijkl}'s and $\beta_{j(i)}$'s are independent .

(4) $\mathrm{Var}(y_{ijkl}) = \sigma^2_\beta + \sigma^2_\varepsilon$.

where, subjects nested within sequences is random, and all other factors are fixed .

Hypotheses of the model :

(1) $H_{01} : \alpha_i = 0$, i. e., testing no significant sequence effect .

(2) $H_{02} : \sigma^2_\beta = 0$, i. e., testing no significance difference between subjects nested within sequences .

(3) $H_{03} : \gamma_i = 0$, i.e., testing no significant sequence (group) effect .

(4) $H_{04} : \delta_i = 0$, i.e., testing no significant formulation (treatment) effect .

The statistical model in (1) can be re-written in matrix notations as follows :

$$\underline{Y} = \mu \underline{1} + \underline{X}\beta + \underline{Z}\theta + \underline{e} \qquad .....(2)$$

Where, $\underline{Y}$ is (nx1) response vector , $\underline{1}$ is (nx1) vector of ones, $\underline{X}$ is (nx6) incidence matrix of the fixed effects factors in $\underline{\beta}$, $\underline{\beta} = (\alpha_1, \alpha_2 , \gamma_1 , \gamma_2 , \delta_1 , \delta_2)'$, $\underline{Z}$ is (nx2) matrix of the random effects factor $\underline{\theta}$, $\underline{\theta} = (\theta_1 , \theta_2)'$ and $\underline{e}$ is a (nx1) vector of errors .

4.2 The Multivariate Analysis of Variance MANOVA :

The proposed approach suggests the use of the MANOVA method to test for statistical bioequivalence which will have an advantage over the ANOVA method, where, the MANOVA takes into consideration the joint distribution between all dependent variables C_{max} , T_{max} , AUC_{0-t} , $AUC_{t-\infty}$. Using the same analogy of the

ANOVA model in (2), the expected value of the MANOVA model can be written as follows :

$$E(\underline{Y}) = \underline{1}\mu + \underline{X}\Psi \qquad .....(3)$$

where,

$\underline{Y} = [\underline{Y}_1 \mid \underline{Y}_2 \mid \underline{Y}_3 \mid \underline{Y}_4]$ is (n x 4) matrix, because we have four dependent variables C_{max}, T_{max}, AUC_{0-t}, $AUC_{t-\infty}$. While, $\underline{Y}_1$, $\underline{Y}_2$, $\underline{Y}_3$ and $\underline{Y}_4$ are (n x 1) vectors, represent the dependent variables C_{max}, T_{max}, AUC_{0-t} and $AUC_{t-\infty}$ respectively with vector of means $\underline{\mu} = [\mu_1 \ \mu_2 \ \mu_3 \ \mu_4]$.

$$\underline{\Psi} = \begin{vmatrix} \alpha_{11} & \alpha_{21} & \alpha_{31} & \alpha_{41} \\ \alpha_{12} & \alpha_{22} & \alpha_{32} & \alpha_{42} \\ \gamma_{11} & \gamma_{21} & \gamma_{31} & \gamma_{41} \\ \gamma_{12} & \gamma_{22} & \gamma_{32} & \gamma_{42} \\ \delta_{11} & \delta_{21} & \delta_{31} & \delta_{41} \\ \delta_{12} & \delta_{22} & \delta_{32} & \delta_{42} \end{vmatrix}$$

Each of the four models $\underline{Y}_1$, $\underline{Y}_2$, $\underline{Y}_3$, $\underline{Y}_4$ can be estimated separately, which will give the same result as in the univariate analysis . However, if we want to take into consideration the joint distribution of the dependent variables, there are (n x 4) errors that are independent across observations, but not across dependent variables .

Hypotheses of the model :

(1) H_{01} : No overall significant sequence effect .

(2) H_{02} : No overall significance difference between subjects nested within sequences .

(3) H_{03} : No overall significant sequence (group) effect .

(4) H_{04} : No overall significant formulation effect .

For estimation purpose the model in (3) can be re-written as follows :

$$E(\underline{Y}) = \underline{W}\,\underline{B} \qquad .....(4)$$

Where $\underline{W} = [\underline{1} \mid \underline{X}]$, $\underline{B} = [\underline{\mu} \mid \underline{\Psi}]$. The multivariate general linear hypothesis can be written as : $H_0 : \underline{LBM} = \underline{0}$, where $\underline{M}$ is the (4 x 4) identity matrix and $\underline{L}$ is the (1 x 4) linear contrast . The multivariate test first construct the matrices $\underline{H}$ and $\underline{E}$ that correspond to the numerator and denominator of a univariate F-test as in SAS User Guide,1989, as follows :

$$\underline{H} = \underline{M}' (\underline{L}\hat{\underline{B}})' (\underline{L}(\underline{W}'\underline{W})^{-1} \underline{L}')^{-1} (\underline{L}\hat{\underline{B}}) \underline{M}$$

$$\underline{E} = \underline{M}' (\underline{Y}'\underline{Y} - \hat{\underline{B}}' (\underline{W}'\underline{W}) \hat{\underline{B}}) \underline{M}$$

Four tests are provided by SAS program :

1. Wilks' Lambda = $\det(\underline{E})/\det(\underline{H}+\underline{E})$
2. Pillai's trace = $\text{trace}(\underline{H}(\underline{H}+\underline{E})^{-1})$
3. Hotelling-Lawley trace = $\text{trace}(\underline{E}^{-1}\underline{H})$
4. Roy's maximum root = λ , largest eigenvalue of $(\underline{E}^{-1}\underline{H})$.

All four tests are exact F-tests, and the only approximation F-test is for testing the random overall significance difference between subjects nested within sequences .

5. NUMERICAL EXAMPLE :

To introduce the proposed approach clearly to statisticians who desire to work in bioequivalence and to users interested in this subject, a numerical example of real data in APPENDIX-A, is illustrated with SAS computer programs in APPENDIX-B . Data was obtained by permission from Center of Pharmaceutical Bioavailability, Faculty of Pharmacy, Tanta University .

5.1 Design of the Experiment :

The study was conducted with 24 healthy male volunteers, according to a two-way crossover design with one week wash-out period . The goal of the study is to provide statistical evidence of bioequivalence between a reference drug called Naprosyn produced by " Syntex Co.", a foreign company, and a test drug called Naprofen produced by "Nile Co. for Pharmaceutical Drugs", an Egyptian Company . The pharmacokinetic parameters obtained for this study are : C_{max}, T_{max}, AUC(0-48hr) and AUC(48-∞) .

TABLE-1

DESIGN OF THE EXPERIMENT

Sequence	Subjects	Period I	Period II
1	1, 2, 3, 4, 5, 6, 7, 8, 9, 10 , 11, 12	R = 1	T = 2
2	13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24	T = 2	R = 1

T : Test Drug (Naprofen), R : Reference Drug (Naprosyn) .

TABLE-1 shows that this experiment was done by a random selection of 12 volunteers (subjects) numbered 1,2,3,...,12 and they were considered the 1st group

(sequence 1), while the other 12 volunteers were given the remaining numbers 11,12,13,...,24 and they were considered the 2nd group (sequence 2) . In period one, volunteers of the 1st group were given the reference drug (Naprosyn) and the 2nd group were given the test drug (Naprofen) . Plasma levels of the chemical substance called "Naproxen" were measured by taking samples of blood from every volunteer during the following times zero, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4, 6, 8, 10, 12, 24, 48 . Then after one week-wash out, in period two, drugs were switched between volunteers as shown in TABLE-1 .

According to this design, the incidence matrix $\underline{X}$ will be a (48 x 6) matrix and can be written as follows :

$$\underline{X} = \begin{vmatrix} 1_{12} & 0_{12} & 1_{12} & 0_{12} & 1_{12} & 0_{12} \\ 0_{12} & 1_{12} & 1_{12} & 0_{12} & 0_{12} & 1_{12} \\ 1_{12} & 0_{12} & 0_{12} & 1_{12} & 0_{12} & 1_{12} \\ 0_{12} & 1_{12} & 0_{12} & 1_{12} & 1_{12} & 0_{12} \end{vmatrix}$$

5.2 The statistical findings :

5.2.1 The Univariate Analysis of Variance ANOVA :

The Statistical Analysis System SAS Ver. 6.10 on IBM PC, was used to perform a (2X2) mixed effects ANOVA model in (1) on the natural logarithms of all dependent variables . All main effects were tested against the mean square error MSE term except, the sequence effect which was tested against the mean square term for subjects within sequence because of the randomness nature of the subjects . Results are shown in TABLES-2, 3, 4 and 5 .

Class Level Information

Class	Levels	Values
SEQ	2	1 2
SUBJ	24	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
PER	2	1 2
TMT	2	1 2

For the natural log of C_{max}, results in TABLE-2 shows that the only significant difference is the difference between subjects within sequences, where the probability of the F-test is 0.003, which indicates a high significance difference between

volunteers (subjects) in each group (sequence) . While there is no significant difference between sequences (Pr > F is 0.8197) . It worth noting that, for testing sequence effect, the mean square error of subjects within sequences should be considered and not the MSE of the whole model . Also, there is no significant difference between time periods (Pr > F is 0.1605), and there is no significant difference between treatments (test and reference drug), (Pr > F is 0.2345) .

TABLE-2

(1) Dependent Variable: $\ln(C_{max})$

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	25	0.79059252	0.03162370	4.21	0.0006
Error	22	0.16530653	0.00751393		
Corrected Total	47	0.95589905			

Source	DF	Type I SS	Mean Square	F Value	Pr > F
SEQ	1	0.00184314	0.00184314	0.25	0.6253
SUBJ(SEQ)	22	0.76166818	0.03462128	4.61	0.0003
PER	1	0.01585276	0.01585276	2.11	0.1605
TMT	1	0.01122845	0.01122845	1.49	0.2345

Tests of Hypotheses using the Type III MS for SUBJ(SEQ) as an error term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SEQ	1	0.00184314	0.00184314	0.05	0.8197

TABLE-3 represents the results of ANOVA applied to $\ln(T_{MAX})$. It is clear that we have about the same results, where the only significant difference is the difference between subjects within sequences, (Pr > F is 0.0199) which indicates a significance difference between volunteers (subjects) in each group (sequence) . While there is no significant difference between the other factors, considering 5% significance level .

TABLE-3

(2)Dependent Variable: $\ln(T_{max})$

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	25	1.48129673	0.05925187	2.26	0.0284
Error	22	0.57580679	0.02617304		
Corrected Total	47	2.05710351			

Source	DF	Type I SS	Mean Square	F Value	Pr > F
SEQ	1	0.01370016	0.01370016	0.52	0.4770
SUBJ(SEQ)	22	1.41706023	0.06441183	2.46	0.0199
PER	1	0.02879110	0.02879110	1.10	0.3056
TMT	1	0.02174523	0.02174523	0.83	0.3719

Tests of Hypotheses using the Type III MS for SUBJ(SEQ) as an error term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SEQ	1	0.01370016	0.01370016	0.21	0.6492

TABLE-4 includes the results of ANOVA applied to $\ln(AUC_{0-48})$. We have again about the same results, where the only significant difference is the difference between subjects nested in sequences, (Pr > F is 0.0004), while there is no significant difference between the other factors at 5% significance level.

TABLE-4

(3)Dependent Variable: $\ln(AUC_{0-48})$

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	25	1.39381733	0.05575269	4.42	0.0004
Error	22	0.27743243	0.01261057		
Corrected Total	47	1.67124976			

Source	DF	Type I SS	Mean Square	F Value	Pr > F
SEQ	1	0.04207381	0.04207381	3.34	0.0814
SUBJ(SEQ)	22	1.26679873	0.05758176	4.57	0.0004
PER	1	0.03959057	0.03959057	3.14	0.0903
TMT	1	0.04535422	0.04535422	3.60	0.0711

Tests of Hypotheses using the Type III MS for SUBJ(SEQ) as an error term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SEQ	1	0.04207381	0.04207381	0.73	0.4019

ANOVA applied to $\ln(AUC_{0-\infty})$ is shown in TABLE-5 .Similar results are obtained, there is no significant difference between all factors at 5% significance level, except for subjects nested in sequences, (Pr > F is 0.0063) .

TABLE-5

(4)Dependent Variable: $\ln(AUC_{48-\infty})$

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	25	3.28558114	0.13142325	2.87	0.0074
Error	22	1.00569399	0.04571336		
Corrected Total	47	4.29127513			

Source	DF	Type I SS	Mean Square	F Value	Pr > F
SEQ	1	0.05876408	0.05876408	1.29	0.2691
SUBJ(SEQ)	22	3.03049235	0.13774965	3.01	0.0063
PER	1	0.19378046	0.19378046	4.24	0.0515
TMT	1	0.00254425	0.00254425	0.06	0.8157

Tests of Hypotheses using the Type III MS for SUBJ(SEQ) as an error term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SEQ	1	0.05876408	0.05876408	0.43	0.5204

The Final conclusion from the univariate analysis : Test and Reference drugs are bioequivalent, since there is no significant difference between sequences, periods, and treatments for each of the four dependent variables C_{max}, T_{max}, $AUC_{(0-48hr)}$ and $AUC_{(48-\infty)}$. However, this conclusion is not quiet right yet ! . That is because results were obtained separately for every dependent variable neglecting the intercorrelation between the dependent variables themselves, especially variables such as $AUC_{(0-48hr)}$

and $AUC_{(48-\infty)}$ could be highly correlated . The multivariate approach overcomes this problem .

5.2.2 The Multivariate Analysis of Variance MANOVA :

SAS Packages is used to estimate the MANOVA model in (3), to test for statistical bioequivalence . The following are some of the important results obtained from SAS output :

TABLE-6

Manova Test Criteria and Exact F Statistics
For the Hypothesis of no Overall Sequence Effect
H = Type III SS&CP Matrix for SEQ E = Type III SS&CP Matrix for SUBJ(SEQ)

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilks' Lambda	0.92108337	0.4070	4	19	0.8013
Pillai's Trace	0.07891663	0.4070	4	19	0.8013
Hotelling-Lawley Trace	0.08567805	0.4070	4	19	0.8013
Roy's Greatest Root	0.08567805	0.4070	4	19	0.8013

Results in TABLE-6 provide sufficient evidence that there is no overall significant sequence effect, where the probability of the F-test for all of the four statistics is greater than 5% .

As it was expected before that there exists a significant correlation between the two dependent variables, AUC from time 0 to time 48 and AUC from time 48 to infinity . TABLE-7 contains the partial correlation coefficients of all dependent variables .

TABLE-7

Partial Correlation Coefficients
from the Error SS&CP Matrix / Prob > |r|

DF = 22	LCMAX	LTMAX	LAUC0-48	LAUC48-∞
LCMAX	1.000000	-0.318351	0.305884	0.077499
	0.0001	0.1388	0.1558	0.7252
LTMAX	-0.318351	1.000000	-0.204284	0.366119
	0.1388	0.0001	0.3498	0.0858
LAUC0-48	0.305884	-0.204284	1.000000	0.609604
	0.1558	0.3498	0.0001	0.0020
LAUC48-∞	0.077499	0.366119	0.609604	1.000000
	0.7252	0.0858	0.0020	0.0001

The partial correlation between AUC_{0-48} and $AUC_{48-\infty}$ with other variables remain fixed is about 61% and it is significant at less than 1% level of significance . This result support the idea of using MANOVA beside ANOVA in bioequivalence studies.

TABLE-8 provides statistical evidence of overall significance subjects within sequence effect . Values of the four statistics, Wilks' Lambda, Pillai's Trace, Hotellin-Lawley Trace and Roy's Greatest Root are different, but they still prove that the test is highly significant, which means the there exists an overall significant differences between volunteers . It worth noting also, that the test in this case is not exact but approximate, because subjects within sequence is a random variable .

TABLE-8

Manova Test Criteria and F Approximations
For the Hypothesis of no Overall SUBJ(SEQ) Effect
H = Type III SS&CP Matrix for SUBJ(SEQ) E = Error SS&CP Matrix

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilks' Lambda	0.00584129	2.3558	88	77.61	0.0001
Pillai's Trace	2.69778787	2.0717	88	88	0.0004
Hotelling-Lawley Trace	12.88049509	2.5615	88	70	0.0001
Roy's Greatest Root	5.96028507	5.9603	22	22	0.0001

Results in TABLE-9 shows that there is no overall significant period effect, where the probability of the F-test for all of the four statistics is 0.2199 which is greater than 5% .

TABLE-9

Manova Test Criteria and Exact F Statistics
For the Hypothesis of no Overall PER Effect
H = Type III SS&CP Matrix for PER E = Error SS&CP Matrix

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilks' Lambda	0.75021414	1.5815	4	19	0.2199
Pillai's Trace	0.24978586	1.5815	4	19	0.2199
Hotelling-Lawley Trace	0.33295276	1.5815	4	19	0.2199
Roy's Greatest Root	0.33295276	1.5815	4	19	0.2199

Results in TABLE-10 does not provide strong evidence of rejecting the null hypothesis of no overall treatment effect because the probability of the four tests is 0.0499 which is about 5% .

TABLE-10

Manova Test Criteria and Exact F Statistics
For the Hypothesis of no Overall TMT Effect
H = Type III SS&CP Matrix for TMT E = Error SS&CP Matrix

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilks' Lambda	0.62114461	2.8972	4	19	0.0499
Pillai's Trace	0.37885539	2.8972	4	19	0.0499
Hotelling-Lawley Trace	0.60993105	2.8972	4	19	0.0499
Roy's Greatest Root	0.60993105	2.8972	4	19	0.0499

This result in TABLE-10 shows that the MANOVA techniques was able to detect some differences between reference and test drug, while in contrast in the univariate case using the ANOVA techniques, there were no significant difference between reference and test drug when testing every dependent variable separately . However, we can still reject the hypothesis of no overall treatment effect if we use 10% significance level, or it might be better if we were able to increase sample size .

6. CONCLUSIONS AND RECOMMENDATIONS

Bioequivalence studies have received much attention in recent literatures . For example, Center for Drug Evaluation and Research (CDER) in USA, has published a Guidance for industry in 1992, and updated this guidance in 1997 and lately in 1999 . This guidance provides rules and tests for bioequivalence . Some other researchers wrote about the same subject such as, Schuirmann, D.J. (1987), Liu, Jen-Pei, (1998), Wang, Weizhen, (1999), Friesen, M. and Walker, S., (1999) and others . The major factor of literature about bioequivalence is that, the proper statistical analysis of a standard comparative bioavailability study requires univariate methods only . This paper, suggested the use of multivariate methods as well as univariate methods . That is because, in a two treatment, two-period crossover design, it is not sufficient to look at the direct treatment effects only using the ANOVA technique for every dependent variable (C_{max}, T_{max}, AUC_{0-t} and $AUC_{t-\infty}$) on a separate basis neglecting the inter-correlation between these variables . It is also important to use the MANOVA technique which takes advantage of the data covariance structure to simultaneously test the equality of means from different dependent variables .

A numerical example is provided to illustrate the basic idea of this paper . This example shows that there might be a possibility of detecting non-bioequivalence between reference and test drugs when applying the multivariate techniques, while in contrast the univariate analysis had proved that the two drugs are statistically bioequivalent .

Finally, as it has been seen in this paper that bioequivalence requires very advanced statistical work , especially work with the multivariate analysis . This situation leads us to recommend that, bioavailability centers in Egypt should consult a professional statistician to obtain adequate and proper statistical results .

REFERENCES

CDER, (1999), " ***Guidance for Industry,*** " Center for Drug Evaluation and Research (CDER), U.S. Department of Health and Human Services, Food and Drug Administration (FDA), USA .

Friesen, M. and Walker, S. (1999), "Are the current bioequivalence standards sufficient for the acceptance of narrow therapeutic index drugs? Utilization of a computer simulated warfarin bioequivalence model," ***J. Pharmaceutical Science*** , 2 (1) : 15-22 .

Hauschke, D., Steinijans, V.W., Diletti, E. and Burke, M. (1992), "Sample Size Determination for Bioequivalence Assessment Using A Multiplicative Model," ***J. Pharm. Biopharm.,*** 20: 557-561.

Johnson, R. and Wichern, D. (1982)," ***Applied Multivariate Statistical Analysis,***" N. J. : Prentice Hall .

Khuri, A., Mathew, T. and Sinha, B. (1998), ***" Statistical Tests for Mixed Linear Models, "*** New York : John Wiley & Sons .

Liu, Jen-Pei (1998), "Statistical Evaluation of Individual Bioequivalence," ***Communications in Statistics, Theory and Methods, V27,*** 6, PP 1433-1451 .

Marascuilo, L., and Levin, J. (1983), " ***Multivariate Statistics In the Social Sciences : A Researcher's Guide,***" California : Brooks Cole Publishing Co.

Montgomery, D. C. (1984), " ***Design and Analysis of Experiments,*** " New York : John Wiley & Sons .

Neter, J. and Wasserman, W. (1974), " ***Applied Linear Statistical Models,***" Illinois : Richard D. Irwin, Inc.

SAS Institute, (1989), "***SAS Users Guide : Statistics,***" North Carolina : SAS Institute Inc.

Schuirmann, D.J. (1987), "A Comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability," ***J. Pharm. Biopharm.***, 15: 657-680.

Searle, S. R. (1971), " ***Linear Models,*** " New York : John Wiley & Sons .

Steinijans, V.W., Hauck, W.W., Diletti, E., Hauschke, D. and Anderson, S. (1992), "Effect of Changing the Bioequivalence Range From (0.80, 1.20) to (0.80, 1.25) on the Power and Sample Size," ***J. Clin. Pharm., Therap. and Toxicol.,*** 30: 571-575.

Wang, Weizhen (1999), "On Testing of Individual Bioequivalence," ***Journal of the American Statistical Association,*** Vol. 94, No. 447, 880-887 .

APPENDIX-A

Pharmacokinetic Parameters
Derived After a Single Dose Administration of Reference and Test Drug
To 24 Healthy Male Volunteers

Volunteer No.	Reference Drug (Naprosyn) Syntex Co.				Test Drug (Naprofen) Nile Co.			
	C_{max} (µg/ml)	T_{max} (hr)	$AUC_{0\rightarrow48}$ (ng.hr/ml)	$AUC_{48\rightarrow\infty}$ (ng.hr/ml)	C_{max} (µg/ml)	T_{max} (hr)	$AUC_{0\rightarrow48}$ (ng.hr/ml)	$AUC_{48\rightarrow\infty}$ (ng.hr/ml)
1	80.58	2.5	818.75	69.66	69.26	2.5	840.94	61.44
2	83.58	2.0	1075.12	116.53	81.28	2.5	927.90	79.78
3	67.76	3.0	984.27	62.43	78.42	3.0	1172.23	98.84
4	80.39	1.5	892.15	57.96	74.13	2.0	974.86	106.49
5	72.32	1.5	955.52	61.74	82.22	3.0	923.27	97.19
6	73.36	2.5	908.58	89.17	81.06	3.0	1220.94	127.43
7	69.51	4.0	1097.16	139.65	72.43	3.5	1278.89	151.19
8	83.64	2.5	1059.85	112.50	84.61	3.0	928.55	84.96
9	62.39	3.0	974.41	90.45	76.9	3.0	1457.74	127.75
10	62.34	2.5	807.62	76.36	81.14	2.5	1169.48	106.43
11	81.18	3.0	1098.31	110.65	75.58	3.0	1340.56	147.92
12	72.27	3.5	1009.20	107.53	91.46	2.5	1039.50	69.20
13	72.53	2.5	1014.13	104.14	66.75	3.5	922.16	104.30
14	63.79	3.5	847.0	82.14	70.33	2.5	872.05	57.80
15	75.03	3.0	1021.68	102.88	60.59	3.0	798.47	62.20
16	115.8	1.5	972.03	86.09	105.31	2.0	1142.22	69.20
17	95.71	2.5	1081.05	89.46	93.34	2.5	1179.80	107.68
18	64.52	3.0	939.99	97.74	62.86	3.0	1060.23	100.60
19	62.07	3.0	488.50	40.09	58.51	2.5	568.40	40.63
20	88.94	3.0	1063.82	107.07	85.77	3.0	1237.35	134.70
21	65.38	3.0	1032.59	96.22	79.93	2.5	956.45	76.98
22	71.95	3.0	1186.32	135.65	75.01	2.5	1043.62	90.08
23	70.71	2.5	1075.19	98.81	72.34	3.0	987.63	96.71
24	75.51	3.0	1123.04	112.94	83.97	3.0	1099.57	104.06

APPENDIX-B

SAS Program for ANOVA Model

```
DATA NAPROFEN;
INFILE 'C:\NAPROF.DAT' ;
INPUT SEQ SUBJ PER TMT CMAX TMAX AUC0 AUC48 ;
        LCMAX=LOG(CMAX); LTMAX=LOG(TMAX);
        LAUC0=LOG(AUC0); LAUC48=LOG(AUC48);
    PROC PRINT ;
    PROC GLM ;
          CLASSES SEQ SUBJ PER TMT ;
               MODEL LCMAX = SEQ  SUBJ(SEQ) PER TMT ;
               MODEL LTMAX = SEQ  SUBJ(SEQ) PER TMT ;
               MODEL LAUC0 = SEQ  SUBJ(SEQ) PER TMT ;
               MODEL AUC48 = SEQ  SUBJ(SEQ) PER TMT ;
          RANDOM SUBJ ;
          TEST H=SEQ PER TMT E=SUBJ(SEQ) ;
    RUN;
```

SAS Program for MANOVA Model

```
DATA NAPROFEN;
INFILE 'C:\NAPROF.DAT' ;
INPUT SEQ SUBJ PER TMT CMAX TMAX AUC0 AUC48 ;
        LCMAX=LOG(CMAX); LTMAX=LOG(TMAX);
        LAUC0=LOG(AUC0); LAUC48=LOG(AUC48);
    PROC PRINT ;
    PROC GLM ;
        CLASSES SEQ SUBJ PER TMT ;
            MODEL LCMAX LTMAX LAUC0 LAUC48 = SEQ  SUBJ(SEQ)
            PER TMT ;
        RANDOM SEQ;
        MANOVA H=SEQ  E=SUBJ(SEQ) /PRINTH PRINTE ;
        MANOVA H= SUBJ(SEQ) PER TMT / PRINTE ;
    RUN;
```